AD-A117 558 NAVAL POSTGRADUATE SCHOOL MONTEREY CA F/G 5/9

AN APPLICATION OF DISCRIMINANT ANALYSIS TO THE CONSTRUCTION OF --ETC(U)

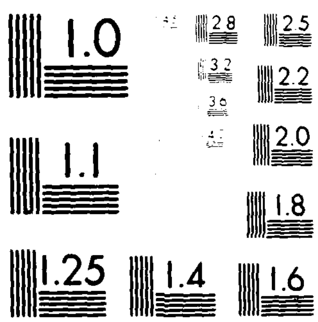
MAR 82 W J HAFFEY

UNCLASSIFIED NL

1 OF 1

AD A

117558



MICROCOPY RESOLUTION TEST CHART

AD A117558

NAVAL POSTGRADUATE SCHOOL
Monterey, California

DTIC
ELECTE
JUL 29 1982
D

THESIS

AN APPLICATION OF DISCRIMINANT ANALYSIS
TO THE CONSTRUCTION OF A
PERFORMANCE INDEX FOR MARINE OFFICERS

by

William Joseph Haffey

March 1982

Thesis Advisor: H. J. Larson

DTIC FILE COPY

Approved for public release; distribution unlimited.

82 07 28 031

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A117568	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Application of Discriminant Analysis to the Construction of a Performance Index for Marine Officers		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; March 1982
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William Joseph Haffey		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1982
		13. NUMBER OF PAGES 95
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Discriminant Analysis
Performance Evaluation
Fitness Report

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study consists of a discriminant analysis of composite marks for 23 selected categories in section B of the Marine Corps fitness report. The data for the study were taken from all the fitness reports on record for those officers in the grades of Captain, Major, and Lieutenant Colonel who appeared before promotion boards during FY81.

Discriminant scores were computed for all officers of

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(20. ABSTRACT Continued)

a particular grade and those officers were then ranked according to this score. It is shown that this ranking closely approximates a ranking by quality and so proves the discriminant score to be a viable performance index, a term whose definition and background are covered in the thesis.

Finally, the unique applicability of the discriminant analysis technique to the performance index problem is demonstrated. While generally unaffected by the distribution of marks within a category, the weight assigned to each category in the discriminant function is very much influenced by the consideration given that category when the promotion decision is made.

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

DTIC COPY INSPECTED 2

Approved for public release; distribution unlimited.

An Application of Discriminant Analysis to the Construction of a Performance Index for Marine Officers

by

William Joseph Haffey
Captain, United States Marine Corps
B.S., The Ohio State University, 1972

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1982

Author: William J. Haffey

Approved by: Thesis Advisor

Second Reader

Chairman, Department of Operations Research

Dean of Information and Policy Sciences

ABSTRACT

This study consists of a discriminant analysis of composite marks for 23 selected categories in section B of the Marine Corps fitness report. The data for the study were taken from all the fitness reports on record for those officers in the grades of Captain, Major, and Lieutenant Colonel who appeared before promotion boards during FY81.

Discriminant scores were computed for all officers of a particular grade and those officers were then ranked according to this score. It is shown that this ranking closely approximates a ranking by quality and so proves the discriminant score to be a viable performance index, a term whose definition and background are covered in the thesis.

Finally, the unique applicability of the discriminant analysis technique to the performance index problem is demonstrated. While generally unaffected by the distribution of marks within a category, the weight assigned to each category in the discriminant function is very much influenced by the consideration given that category when the promotion decision is made.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

In this thesis, a technique is developed for constructing a performance index from the markings on section B of the Marine Corps fitness report through the use of a statistical tool called Discriminant Analysis. The technique is then applied to the markings on fitness reports of 967 officers of the rank Captain through Lieutenant Colonel who appeared before promotion boards during the 1981 fiscal year.

The idea of a performance index is explored, criteria for measuring the effectivenss of such indices are suggested, and uses for the index--both current and envisioned--are studied. In addition, past attempts at constructing a performance index are examined.

The theory underlying the discriminant analysis technique is developed, not rigorously but in sufficient detail to allow readers to appreciate the applicability of the procedure and the significance and meaning of the results.

Finally, the weights assigned to each of the fitness report categories in the discriminant function are examined. The influence of such factors as the distribution of marks within the categories on these weights is studied and the final conclusion lends considerable support to the use of discriminant analysis in the construction of the performance index.

II. BACKGROUND

A. THE MARINE CORPS FITNESS REPORT

The USMC fitness report currently in use is NAVMC Form 10835. It is in 3 sections. The first section, section A, records such administrative data as name, grade, current billet and choices for further duty assignments. Section C provides space for a subjective word picture of the individual reported on. The second section, section B, is shown in Figure 1. It consists of several fixed-choice markings on several categories or dimensions of performance. For the most part, there are 6 markings for each category. These markings are unsatisfactory, below average, above average, excellent and outstanding. There is also a provision for a rating of "not observed" in case a particular performance dimension is not demonstrated during the period of the report. Block 15a has the same 6 fixed-choice markings as the first group with an additional block between each marking. Block 16 has a different scale of markings, as can be seen in Figure 1.

Six of the blocks in section B will not be considered in the study: blocks 15b, 15c, 17, 18, 19 and 20.

Throughout his career, each Marine officer is rated by means of the fitness report on several occasions. The rating occurs at least every 6 months or sooner if

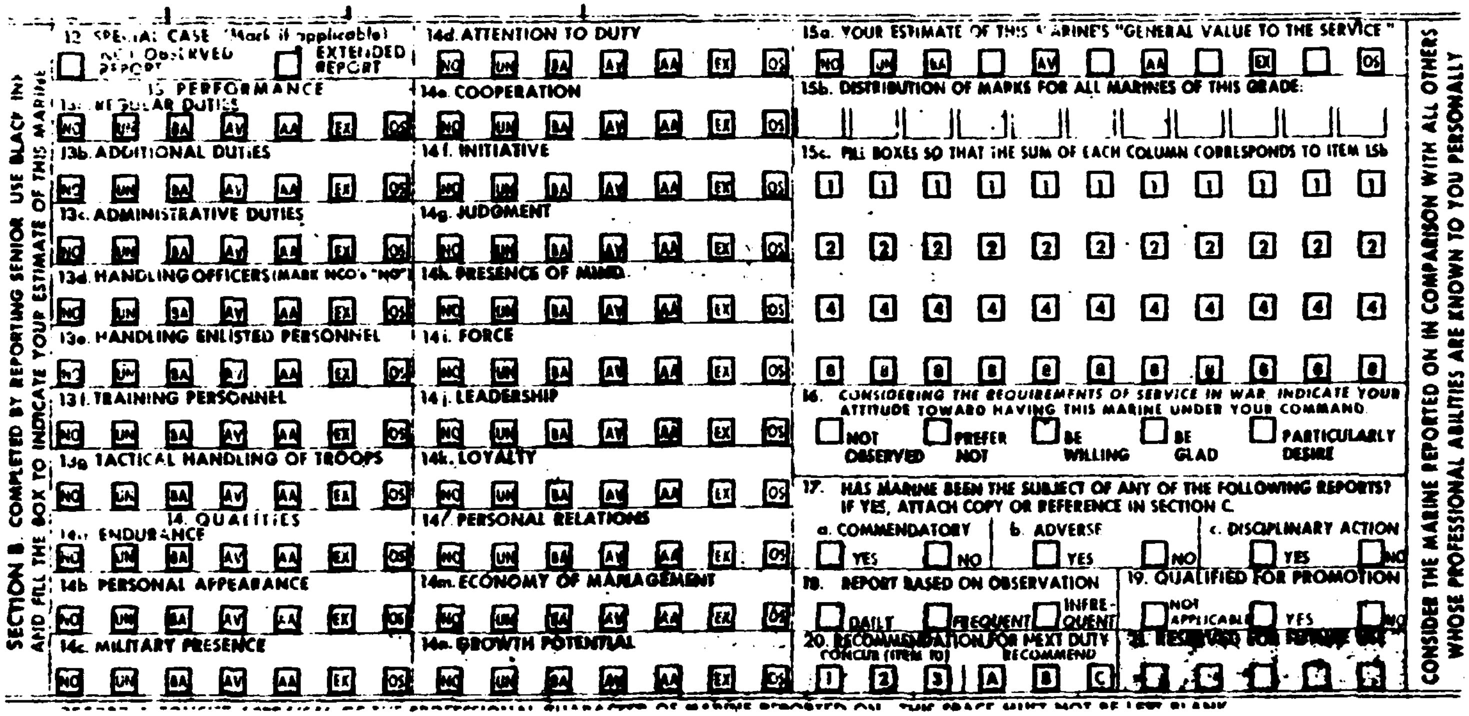


FIGURE 1. SECTION B, MARINE CORPS FITNESS REPORT

occasioned by a promotion, transfer, change of commanding officer, etc. The fitness report is the principle means utilized by promotion boards in judging the qualifications of officers for advancement to the next higher grade.

B. THE PERFORMANCE INDEX

The term performance index, in the context of this thesis, will mean some quantitative, composite measure reflecting an officer's quality or promotability. It would indeed be difficult to separate these two attributes (quality or promotability) to any great extent, for in the final analysis, it is the quality of the officer which (hopefully) results in the quality of his record and thence his advancement. At any rate, as will be seen in later chapters of this thesis, it will be the extent of the similarity of an officer's record to that of a group of promoted officers that will determine the measure of his quality--or his performance index.

The composite nature of the performance index is stressed. Although there is certainly no substitute for the multi-dimensioned performance measure represented by all the categories of section B, together with all the other perhaps non-quantifiable indicators of quality considered by a promotion board, there are certainly circumstances where a single, composite measure of quality would

be useful and indeed, necessary. Some of those circumstances will be highlighted later on.

In comparing and contrasting various performance indices (or indeed any method of performance appraisal) it is necessary to establish and define the criteria to be used in determining the effectiveness or worth of each index or method.

First, and perhaps foremost among the criteria that could be employed is the validity of the index in reflecting the true quality of the officer. It is very true that any variable describing an officer could be used as a performance index--for example, hair color or date of birth--but it's doubtful that many of these types of indices could be described as valid. In practice, when judging the validity of a proposed index, the measurement of validity will be both intuitive and, possibly, quantitative. The intuitive judgement comes into play in rejecting the vast majority of possible indices (such as the two just mentioned).

The quantitative measure of validity, when applied, will be in two phases. The first phase involves a measure that was actually employed at Headquarters, U.S. Marine Corps in the early months of this project. The measurement was made as follows:

The proposed performance index was computed for all the officers whose records appeared before a particular promotion board. The officers were then ordered from highest to

lowest according to this computed index. The first n officers on this list, where n was the actual number of officers who were promoted by the board, were chosen as a sample. The criterion measure was the sample proportion (p) of promoted officers. Although perhaps not a theoretically sound measure, statistically, several characteristics of this measure were attractive and readily apparent. If the performance index utilized was a "perfect" indicator of promotability/quality, p should be 1.0. If the proposed index was in some sense a "poor" indicator, p should approximate the actual promotion proportion for all the officers who were considered. And a computed p that was less than the true promotion proportion would indicate that the index was in a sense negatively correlated with promotability. Again, it is admitted that the utility of this measure lies not in its statistical grounding but in its computational ease, intuitive appeal and its ability to quickly and decisively eliminate some very bad candidates for a performance index.

In Chapter 4, other quantitative criteria will be proposed and applied when our interest will not be in a "first-cut" measure designed to isolate good or bad ideas for indices but rather in investigating fine differences among several good indices in the form of discriminant functions.

Perhaps not as vital a criterion as validity, practicality has been de-emphasized as a result of the wide-spread use of high-speed computers in the administration and data-storage of the Marine Corps performance report system. Yet, practicality as a criterion still bears mention.

Generally, the practicality of a performance index is a measure of, first, the accessibility of the information that is involved in the computation of the index and, second, the actual computational ease. In fact, the information used in computing all of the potential performance indices mentioned below was readily available in some sort of computer storage medium. Further, the differences in computational difficulty were practically immeasurable. However, the applicability of practicality as a viable criterion can be appreciated using a simple example of a potentially highly accurate index whose practicality is certainly questionable. Such an index would involve the reduction of the written appraisal in section C of the fitness report to a quantifiable measure. Indeed, the information is accessible but the computation has evident limitations in that attaching a single numerical score to a written evaluation on an individual would be a difficult task, to say the least, and would involve the resolution of such issues as the proper scaling of such scores.

These criteria will now be applied to several indices that have either been tried in the past or envisioned for use in the future.

The idea of a performance index is by no means a new one. A form of index has been in use for some time by enlisted promotion boards in the Marine Corps. This index is a weighted, linear combination of selected markings from section B of the same report described in paragraph (a) above. At a typical promotion board session, the records of the candidates for promotion are distributed among the board members by performance index--to ensure that each member is given a representative sample of the candidate group and that no one member is forced to consider a group of predominantly below or above average candidates. Additionally, the sheer number of records that each board member must consider dictates that he also use this index as a first-cut criterion for promotion. That is, a candidate with a relatively high index score is automatically promoted, with a low score automatically not promoted, and with the cases falling in between being more carefully examined before a final decision is made.

The choice of selected markings and the weights assigned to each marking were determined quite subjectively over the years and represent no real concerted attempt at an optimum selection or an optimum weighting, and therefore no real attempt at validity.

A means of distributing records among promotion board members, such as the process described above for the enlisted board, was one of the major potential uses envisioned for an officer's performance index. Another use of such an index could be in regression/correlation studies. For instance, the performance indices for a group of officers might be regressed on such variables as commissioning source, undergraduate degree, or officer candidate school standing to discover the degree of correlation between each of these factors and the officer's quality--as measured by his performance index. Indeed, once such a regression function was constructed in this manner, an officer candidate's potential success should be well-estimated--certainly a valuable piece of information for a recruiter.

Among the first potential indices considered was a measure commonly known as the "truth teller". It was computed using a complicated formula involving the markings on items 15a, 15b, and 15c on section B of the fitness report. A combination of the number of officers ranked even, the number ranked above, and the number ranked below--over all the fitness reports considered--gave a percentage grade, the truth teller. It was and, in some cases, still is being utilized by some officer selection boards in much the same way as the index for enlisted boards--as a way of

distributing the records by quality among the members of the selection board. Its effectiveness as an index as measured by the criterion of practicality is indisputable. However, its accuracy can only be described as fair. Indeed, recent investigation uncovered several instances of misuse of this measure on the part of officers completing reports on subordinates. As a result, in too many cases the computed "truth teller" for an officer was an invalid measure of his quality. Not surprisingly, the quantitative criterion--the measure p, as described above--when applied to this particular index yielded disappointing results.

Another potential candidate for a performance index was the score on the military General Classification Test (GCT). The test is administered to every officer upon entry, is a permanent part of his military record, and so fares well with the practicality criterion. Yet, in no case was the computed p on the quantitative accuracy criterion higher than that for the truth teller.

It was then decided to attempt to construct a composite, linear, weighted combination of the marking categories in section B to yield a performance index, much like the system in use by the enlisted promotion board. However, unlike the enlisted system which weights only selected marking categories, this scheme would consider all marking

categories (the only exceptions being those outlined in paragraph B above). The problem of assignment of weights to categories was resolved by the Delphi technique wherein a memo requesting suggested weighting schemes was circulated among several officers across as broad a spectrum of grades and expertise as possible. From these suggested schemes, a final vector of category weights was fashioned. The effectiveness of this performance index as measured by the accuracy criterion p proved that this approach (the weighted linear combination) was a good one. The measured p for all three promotion boards was consistently between .88 and .90.

Yet there was still a basic flaw in both this scheme and to an even greater extent the one employed by the enlisted promotion boards. As was mentioned, the objective of the performance index was to present a composite score indicating the quality or promotability of a particular officer. The schemes presently in use represented the officer's relative standing among his peers when the criterion used might not necessarily represent his promotability in terms of what an actual promotion board action would reflect. Rather, these performance index scores reflected only the officer's quality measured against what a few selected questionnaire respondents <u>thought</u> quality should be.

A solution seemed to be present in the fact that at the end of a board, two pieces of information were in hand. First, the exact composition of the two groups (i.e., promoted group and not promoted group) was known. Secondly, a multi-dimensional observation vector (representing the markings in section B) was known for each officer in each group.

Taking advantage of this information, a possible alternative approach was seen to be as follows.

Compute the mean values on each of the marking categories for the promoted group and do the same for the not promoted group. Call this vector of mean values for each group the group centroid. A possible performance index for an individual, then, could be the extent of the similarity of an individual's markings to the promoted group centroid.

This idea is basic to the discriminant analysis technique and will be developed in some detail in the next chapter.

III. DISCRIMINANT ANALYSIS

The following discourse on the discriminant analysis technique is intended primarily to enable the reader to appreciate the applicability of the technique to the construction of a performance index. The mathematics of the technique will be presented in a somewhat elementary fashion so as not to preclude anyone from gaining a full appreciation of the underlying ideas. For a more rigorous discussion, the interested reader is directed to any of the references listed in the bibliography, especially the book by Tatsuoka.

Before beginning, it should also be mentioned that the discriminant analysis technique is applicable to the study of any number of groups greater than or equal to two. The discussion, however, will address primarily two-group analysis since the performance index problem and attendant data are both concerned with two groups--promoted and not promoted officers. In addition, the derivation of the discriminant function is much more straightforward in the case of two groups. In fact, Tatsuoka [Ref. 1] shows that two-group discriminant analysis is in many ways identical to multiple regression analysis. A rigorous discussion of the similarities (in a conceptual sense) between regression analysis and discriminant analysis can be found in Rulon [Ref. 2].

A. THEORY

Discriminant analysis is essentially a procedure for quantifying the differences between groups, albeit special groups. The underlying assumptions concerning each group are as follows:

(1) Each group t (t = 1,2) has a number of members, each member being defined as a p-element vector of observations or measurements,

(2) The groups being investigated are separate and identifiable and

(3) The observations are assumed to have a multivariate normal distribution with equal covariance matrices.

NOTE: further notational conventions that will be utilized throughout this chapter are as follows:

- X_{tij} is defined as the observation on the jth variable for the ith member of group t
- $(X_{ti1}, X_{ti2}, X_{ti3}, \ldots, X_{tip})$ is therefore the p-element vector defining the ith member of group t
- $(\bar{X}_{t\cdot 1}, \bar{X}_{t\cdot 2}, \bar{X}_{t\cdot 3}, \ldots, \bar{X}_{t\cdot p})$ is the p-element vector of observed means on the p variables for group t.

Assumption (3) would appear to be rather restrictive especially in the context of the fitness report data that will be studied, given that many of the p variables will exhibit a highly negatively-skewed distribution. However,

Eisenbeis [Ref. 3] notes that there exists evidence that non-multivariate normal data may be used in a discriminant analysis without significantly biasing the results. (But, as will be seen in the next chapter, the actual vectors of fitness report category marks that are entered into the discriminant analysis are composed of marks that represent means over several reports, and, as such, these marks tend to be more normally distributed than the original data).

The study of the differences between groups entails a study of which particular elements or groups of elements in the p-element vectors define a dimension or direction along which the major group differences occur. This idea of certain elements or groups of elements defining a dimension of greatest difference would logically entail an algebraic notion of linear combinations; specifically, a linear combination of the original p observation variables that will somehow give a large difference between group means on that linear combination relative to the inherent within-group differences among values on that same combination.

This linear combination of the original observation variables is called the discriminant function and the coefficients assigned to each of the original variables are called the discriminant function coefficients. The discriminant function yields a discriminant score for the ith member of group t which is defined as:

$$Y_{ti} = v_1 X_{ti1} + v_2 X_{ti2} + \cdots + v_p X_{tip}$$

where v_i, $i = 1,2,\ldots,p$ are the discriminant function coefficients; $X_{ti1}, X_{ti2}, \ldots, X_{tip}$ are the p original observation variables for the ith member of group t; and Y_{ti} is the discriminant score for that same member. As is evident, once this discriminant function with its coefficients is devised, each member of each group will have associated with it a discriminant score.

Figure 2 illustrates this notion for the special case of $p = 2$. In this Euclidean 2-space, each group member is located by a single dot and the collection of dots representing one group is outlined. In turn, the t-th group centroid is denoted by $(\overline{X}_{t\cdot 1}, \overline{X}_{t\cdot 2})$. The group centroid is simply the 2 element vector representing the mean values of each original observation variable within a group. Within the restricted scope of this treatment, what will be of interest are the empirical distributions of the discriminant scores Y within each group, represented in the figure by $f_1^*(Y)$ and $f_2^*(Y)$. The empirical distributions reflect the projections of the original p observation variables onto the line L, where the projection is in fact the $E^p \rightarrow E^1$ transformation accomplished by the discriminant function. Also of interest is the fact that given a particular orientation of L, the distance between the projections of the group centroids onto L is maximized and the overlap between the 2 groups is minimized. Intuitively, it would seem that the orientation of L should be such that it is parallel to

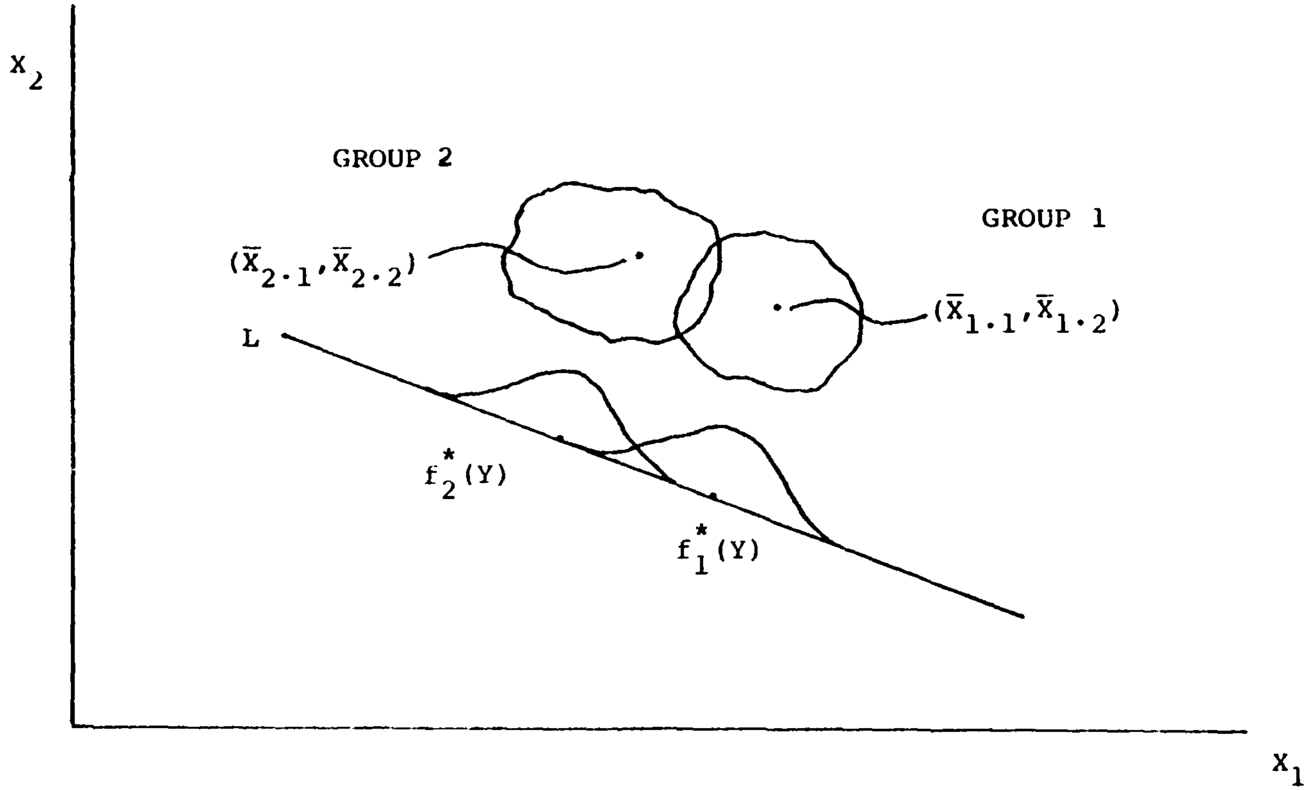


FIGURE 2. GEOMETRIC INTERPRETATION OF DISCRIMINANT ANALYSIS

the line connecting the two group centroids in order that such a maximum centroid separation would occur. In most cases, this is indeed true.

Discriminant analysis is the technique that will determine the optimum discriminant coefficients and thus the optimum orientation of L so that indeed the differences in univariate means of the discriminant scores is maximized relative to the within-group variance in those scores.

B. DERIVATION OF THE DISCRIMINANT FUNCTION COEFFICIENTS

The first step toward determining an optimum linear combination of the set of p variables such that the two group means will differ widely between themselves on that linear combination relative to the within-group differences on that linear combination is to suggest a criterion with which to measure "optimum".

Once a linear combination of the p variables has been constructed, one is dealing with a single transformed variable. Thus, the familiar F-ratio for testing the significance of the difference between two group means on a single variable seems applicable. Indeed, it would seem that the optimum linear combination would yield a single variable that would maximize this F-ratio given by:

$$F = \frac{SS_b/(K-1)}{SS_w/(N-K)} = \frac{SS_b}{SS_w} \cdot \frac{(N-K)}{(K-1)}$$

where N is the total number of members of both groups (n_1 + n_2 in this case), K is the number of groups (K = 2 in this case), SS_b is the sum of the squared deviations of each group mean around the grand mean, and SS_w is the sum of the squared deviations of each group member's score around the group mean. That is,

Y_{ti} is the discriminant score for the ith member of group t,

$\overline{Y}_{t.}$ is the group t discriminant score mean,

$\overline{Y}_{..}$ is the grand discriminant score mean and so,

$$SS_b = \sum_{t=1}^{2} n_t (\overline{Y}_{t.} - \overline{Y}_{..})^2 \qquad (3\text{-}1)$$

and

$$SS_w = \sum_{t=1}^{2} \sum_{i=1}^{n_t} (Y_{ti} - \overline{Y}_{t.})^2 \qquad (3\text{-}2)$$

Since the second factor in the F-ratio, (N-K)/(K-1), is a constant for any given problem (where N and K are fixed), the first factor SS_b/SS_w is the only essential quantity for measuring how widely a set of group means differ among themselves relative to the amount of variability present within the groups. This ratio SS_b/SS_w is what is called the discriminant criterion. Writing this discriminant criterion now in terms of the linear combination of original variables

instead of discriminant scores Y would involve making the following substitutions into equations 3-1 and 3-2:

$$Y_{ti} = v_1X_{ti1} + v_2X_{ti2} + \dots + v_pX_{tip}$$

$$\bar{Y}_{t\cdot} = \frac{1}{n_t}\left(\sum_{i=1}^{n_t} v_1X_{ti1} + \dots + \sum_{i=1}^{n_t} v_pX_{tip}\right)$$

$$\bar{Y}_{\cdot\cdot} = \frac{1}{N}\left(\sum_{t=1}^{2}\sum_{i=1}^{n_t} v_1X_{ti1} + \dots + \sum_{t=1}^{2}\sum_{i=1}^{n_t} v_pX_{tip}\right)$$

Now maximizing the criterion over all vectors $V = (v_1, v_2, \dots, v_p)$ will yield the optimum V which will be the vector of discriminant function coefficients we are seeking.

Before continuing, mention should be made of the technique employed by most commercial discriminant analysis packages--particularly the SPSS routine which was used to perform the bulk of the analysis in this study--in deriving the discriminant function. Essentially, it is identical to the method outlined above with but a few notable exceptions.

Instead of expressing the discriminant function as a linear combination of the original variables $X_{ti1}, \dots, X_{tip}$, as in

$$Y_{ti} = v_1X_{ti1} + v_2X_{ti2} + \dots + v_pX_{tip}$$

the discriminant function is instead expressed as

$$Z_{ti} = v_1 Z_{ti1} + v_2 Z_{ti2} + \cdots + v_p Z_{tip}$$

where Z_{tij}, $j = 1,2,\ldots,p$ represents the standard deviations of the original variable X_{tij} from its group mean (i.e., Z_{tij} represents the standardized score). When these standardized variables are substituted for the original variables in equations 3-1 and 3-2, discriminant function coefficients can be derived to result in discriminant function values (discriminant scores) that are themselves in standard from. This means that, over all cases in the analysis, the discriminant scores will have a mean of 0 and standard deviation of 1. The discriminant score of any one particular individual will represent the *number of standard* deviations that case is away from the mean for all cases on the discriminant function.

Morrison [Ref. 4] outlines several reasons why the standardized values of the original variables are used in the analysis in place of the original variables, one of which is the fact that since our analysis is concerned with the "distance" between two groups, and since statistical distance is normally measured in terms of standard deviations, one is justified in normalizing the original variables.

The standardized discriminant coefficients (those computed using the standardized original variables) yield a great deal of information about the original variables.

When the sign is ignored, each coefficient represents the relative contribution of its associated variable to the discriminant function score. (The interpretation is analogous to the interpretation of beta weights in multiple regression.) The sign merely indicates whether the contribution is positive or negative. One can even say that the greater the magnitude of a variable's coefficient, the more powerful that variable is in discriminating between the two groups.

A question of real interest to the architect of any performance evaluation system is what statistics associated with a particular variable common to the two groups render one more discriminating than the other. As part of the analysis presented in Chapter V, relationships between the variable's coefficient magnitude and such statistics on that variable as within group variance, correlation and range and between-group F statistics will be examined.

C. VARIABLE SELECTION METHODS

Referring again to the first few sentences of this chapter, recall that discriminant analysis requires that each group member be defined by a p-element vector of measurements or observations. It was this vector $(X_{ti1}, X_{ti2}, \ldots, X_{tip})$ that was used to compute the discriminant function and that set of discriminant coefficients $V = (v_1, v_2, \ldots, v_p)$ such that the linear weighted combinations of the original p variables determined by the discriminant coefficients would have the

maximum separation between group means relative to the within-group differences.

In many instances, however, the full set of p variables may contain an excess of identical information about group differences, perhaps because of a high degree of correlation among certain of them. Or perhaps some of the variables may not be very useful in discriminating between the groups. It might therefore be useful to be able to determine a reduced set of q variables ($q < p$) that perform almost as well as the full set p in the discriminant analysis.

Eisenbeis [Ref. 5] and Nie [Ref. 6] outline several such methods for reducing the dimension of the variable vector. Only one will be discussed now and indeed utilized in the next chapter. This particular method is the forward-selection procedure using the Wilks' Lambda statistic criterion.

In a forward-selection procedure, the process begins by selecting the one variable which has the highest value on a particular criterion--in this case, the Wilks' Lambda which will be described in detail below. Next, the remaining p-1 variables are paired one-at-a-time with the already-selected variable and the criterion is computed for each pair. The variable which, when paired with the first one selected, yields the best value on the criterion then becomes the second variable selected. These two variables are then combined with each of the p-2 remaining variables and the criterion then determines the third variable to be entered into the

analysis. This process continues until either all the variables are selected or the inclusion of further variables yields less than a pre-determined improvement in the criterion measure.

At the same time, as variables are selected for inclusion in the analysis, some of the variables previously selected may lose some of their discriminating power primarily because the information they contain might now be available in some other combination of the other variables. For this reason, before each step in the forward selection, all variables currently in the analysis are examined to see if they still make a contribution to discrimination. If not, they are eliminated but are eligible to again enter at a later step.

Once the q variables are selected by this process, the discriminant analysis is performed using this q-vector of observations on each group member.

In this study, the criterion for the stepwise selection is the Wilks' Lambda statistic. Wilks' Lambda, Λ, is defined as follows:

$$\Lambda = \frac{|W|}{|T|}$$

where W and T are the within groups and total sample sums of squares and cross-products (SSCP) matrices defined by Tatsuoka [Ref. 1], and $|W|$ is the determinant of the matrix W.

To examine several of the properties of Wilks' Λ, it is first helpful to see what Λ reduces to when there is but a single observation on each group member, i.e., $p = 1$.

In the case where $p = 1$, $|W|$ reduces to

$$|W| = \sum_{t=1}^{2} \sum_{i=1}^{n_t} (X_{ti} - \bar{X}_{t\cdot})^2$$

which is what SS_w was originally defined as (except that Y_{ti} was substituted for the X_{ti}) and $|T|$ reduces to

$$|T| = SS_w + \underbrace{\sum_{t=1}^{2} n_t(\bar{X}_{t\cdot} - \bar{X}_{\cdot\cdot})^2}_{SS_b}$$

where

- there are $t = 2$ groups, of sizes n_t,
- the ith member of group t is defined by X_{ti} (since there is only one variable characterizing each group member) and
- $\bar{X}_{t\cdot} = \frac{1}{n_t} \sum_{i=1}^{n_t} X_{ti}$, $\bar{X}_{\cdot\cdot} = \frac{1}{N} \sum_{t=1}^{2} \sum_{i=1}^{n_t} X_{ti}$

Therefore,

$$\Lambda(p = 1) = \frac{SS_w}{SS_w + SS_b} = \frac{1}{1 + (SS_b/SS_w)} \qquad (3\text{-}3)$$

But note that the customary univariate F-ratio for a simple analysis of variance (k-group) is given by

$$\frac{SS_b}{SS_w} \cdot \frac{(N-K)}{(K-1)}$$

where $N = n_1 + n_2$, and in this case, $k = 2$. Consequently,

$$\frac{SS_b}{SS_w} = \frac{K-1}{N-K} \cdot F$$

and, when substituting this expression for SS_b/SS_w into 3-3, we obtain

$$\Lambda(p = 1) = \frac{1}{1 + [(K-1)/(N-K)]F} \qquad (3-4)$$

From (3-4) we can see that, at least in the univariate case, there is an inverse relationship between Λ and F. In other words, the larger the disparity among several group means, relative to within-group variability, the larger is F, but the smaller is Λ. It is interesting and useful to note that, according to Tatsuoka [Ref. 1], this relationship also holds in the multivariate case (i.e., $p > 1$).

At each step of the forward selection process using Wilks' Λ as the criterion, that variable which, when combined with the variables already in the analysis, yields the lowest value of Λ will be selected for inclusion. The process will continue until the inclusion of further variables

yields a decrease in Λ which is less than a predetermined value (chosen to be .001 in the analysis). These q variables will then be entered into the discriminant analysis routine.

D. APPLICABILITY TO THE PERFORMANCE INDEX PROBLEM

Refer, again, to Figure 1, the line L and the projections of the original p-element vector of variables onto this line L--the projections being accomplished by the $E^p \rightarrow E^1$ transformation via the discriminant function. Recall that this transformation yields a value called the discriminant score for each member of each group. The empirical distributions of those discriminant scores for the two groups were pictured as $f_1^*(Y)$ and $f_2^*(Y)$.

Refer now to Figure 3(a). The hypothetical empirical distributions of the discriminant scores for the officers who were selected (S) and not selected (NS) are pictured. Also pictured are the locations of the discriminant scores for four hypothetical officers (p_1, p_2, p_3, p_4) whose records were not necessarily used in the analysis but whose military rank is identical to those officers whose records were. Now if we assume that the range of the discriminant scores was (-2,2), Figure 3(b) represents the individual discriminant scores for each member of the two groups against a scale to the right. The mean score for each group is again denoted (as it is in Figure 3(a)) by $\overline{Y}$. Also pictured are the relative locations of the discriminant scores for the four officers.

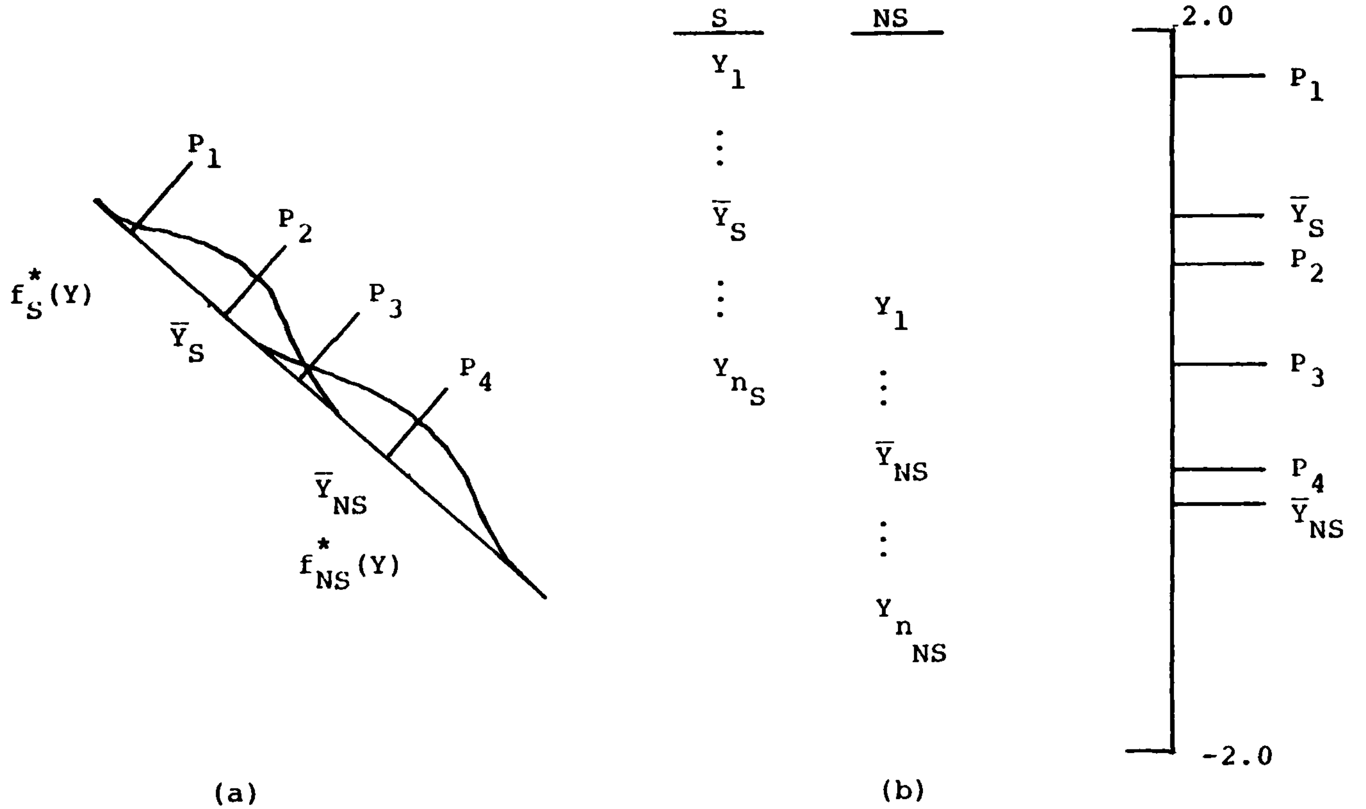


FIGURE 3. CLASSIFICATION OF FRESH CASES AND RANKING BY DISCRIMINANT SCORE

Recall that the transformation of the original p-vector of variables for each member has been accomplished such that the resultant discriminant scores show the greatest disparity between group means relative to within-group differences of those scores.

One of the principal uses for discriminant analysis is in classifying cases that were not previously used in the analysis (hereinafter called fresh cases) into one of the k groups with a particular certainty of having made a correct classification. The classification is made, in a sense, on the basis of the distance or similarity of the fresh case's vector of measurements to the group centroids of each of the k groups. Classification theory is a science in itself and is treated in some detail in several of the listed references. However, one is not too far wrong in saying that a fresh case should be classified as belonging to group k if the distance between the discriminant score of the fresh case and the mean discriminant score for group k is less than the distance between the discriminant score for the fresh case and any other group mean score. For instance, it would appear that the officers labeled p_1 and p_2 should be classified as belonging to the selected group whereas the officer labeled p_4 would more likely belong to the not-selected group.

However, the intent of this thesis is to take this idea one step further and claim that, based on the computed discriminant scores, the record of officer p_1 is in some sense

better than the record of p_2 which in turn is better than p_3 which again is better than p_4. This ordinal ranking can be rationalized in a sense because the discriminant score for officer p_1 is "further away" from the mean discriminant score for the not selected officers than is the score for officer p_2. At the same time, the score for officer p_4 is "closer to" the mean score for the not selected officers than is the score for officer p_3, and so forth.

In fact, the hypothesis to be examined is that a ranking of the discriminant scores of officers of a particular grade from highest score to lowest is a legitimate method of ranking the quality of those same officers. Indeed, an officer's discriminant score would be his performance index.

It is of course recognized that such a ranking by quality has been based indirectly on factors that make up only a portion of the entire record on each officer. Totally ignored are such important performance dimensions as the type of duty performed during the period of a report, the written description on each officer contained in the report and other non-quantifiable points that every selection board rightfully takes into account when rendering its decision on each officer.

However, bear in mind the uses to which such a performance index for officers would be put, as outlined in Chapter II. At no time would such an index be used in the actual promotion decision. At no time has a claim been made that a

performance index would be anything more than an "indicator" of quality, not an actual measure. This particular index certainly fares well on the practicality criterion, and bears study as to its performance on a validity criterion, a study that will be undertaken in the next chapter.

IV. ANALYSIS AND RESULTS

A. DATA

The raw data set for the study was obtained from Headquarters, Marine Corps personnel files. The set consisted of all fitness reports on file for each officer who appeared before the Major, Lieutenant Colonel or Colonel promotion boards during the 1981 fiscal year. The information on file for each report was complete with the exception of the written appraisal in Section C. The name on each report was erased and the SSN encrypted to ensure anonymity. In all, there were 42,314 reports on 967 officers.

The record on each fitness report consisted of the marks on the 23 performance categories in Section B of the report. However, instead of the markings that actually appear on the report (such as OS, EX, etc.), the marks on the record were encoded numerically as shown in Table I.

The number of reports on each officer ranged from 25 to 60 depending, of course, on his rank. To have entered each mark in each report as a separate variable in the discriminant analysis would have been impractical, so it was first decided to construct a single record for each officer. On this single record, the marks in each of the 23 categories would represent a composite of all the marks on that category over all the reports on file. The method for constructing such a single record was an important decision, as well

TABLE I

CODING OF FITNESS REPORT MARKS

Fitness Report Marking	Numerical Equivalent
Blocks 13a to 14n	
OS	9
EX	7
AA	5
AV	3
BA	1
UN	0
Block 15a	
OS	9
EX-OS	8
EX	7
AA-EX	6
AA	5
A-AA	4
A	3
BA-A	2
BA	1
UN	0
Block 16	
Particularly Desire	9
Be Glad	8
Be Willing	7
Prefer Not	6

as a difficult one, for in choosing a scheme for computing a composite mark for each category, it was paramount to describe the tendencies of promotion board members.

Conceivably, there are an infinite number of such schemes but, given resource constraints, only three were devised. It was thought that these three represented logical schemes that would accurately describe most board members behavior when screening an officer's reports.

The first scheme involved computing a simple average (or mean) of the marks on all reports for a particular category as the composite mark for that category. The set of composite category marks computed according to this scheme will be known as the variable set BN.

The second scheme was to constrcut the composite mark for each category as a weighted average over all reports of the marks in that category. The weighting factor for each report's mark in a particular category would be the number of months covered in that report divided by the total number of months covered by all reports in the individual's file. For instance, if the total number of months covered by an individual's reports on file was 120 months, the weighting factor for a report of 12 months duration would be 0.1, and one covering 6 months would be weighted .05. This variable set will be labeled WB.

The third scheme was a variation on the second in that it too involved a weighted average of individual report

marks, but only over those reports written on an individual in his present grade. For instance, in the case of a Major appearing before the Lieutenant Colonels' promotion board, only those reports written on him while in the grade of Major were considered. If all such reports covered a period of 60 months, a report of 12 months duration would have a weighting factor of 0.2. This variable set will be titled GB.

As was mentioned in a previous chapter, there would be considerable appeal in a reduced-dimension set of marks to be employed in computing an officer's discriminant score or performance index. For this reason, an additional subset of each of the original variable sets (BN, WB, and GB) was determined for each officer. The categories to be included in this subset were determined using the forward selection procedure and the Wilks' Lambda criterion described in Chapter III. The subsets determined in this manner will carry the same name as the original set from which the subset was taken, but with a suffixed -W (for example, BNW).

During a preliminary examination of the data, it was found that, in certain particular categories of the fitness report, no marks had been assigned to that category on any of the reports on file (such a missing mark to be hereinafter termed a missing value). As a result, any composite mark computed for that category would necessarily also exhibit a missing value. Unfortunately, packaged discriminant

analysis routines will totally exclude from the analysis the variable set of any officer that is not complete by even one missing value. If, for one of the particular variable sets, a sizeable percentage of the officers exhibited missing values on one or more of the variables in that set, the set would of course be a poor candidate as a possible set of marks to be entered into the discriminant analysis, for two reasons:

First, computing a discriminant function for a group using the variable sets from only a small sample of that group (with no guarantees that the sample is even remotely representative of the whole group) invites questionable results. Second, once such a discriminant function is determined, a valid discriminant score could be computed for only that same small sample of officers since only they would have a complete set of the requisite category marks by which the discriminant coefficients would be multiplied to yield a discriminant score. Herein lay one of the incentives in utilizing the forward-selection procedure to determine a reduced-dimension set of marks. The fewer the marks in the variable set, the smaller the percentage of officers who might exhibit a missing value on any one of the variables in that set.

Conceivably, however, even such a reduced-dimension variable set might still contain variables on which a number of officers exhibit a missing value.

Therefore, it seemed a logical step to construct variable sets composed of variables for which it was known that a vast majority of officers had a valid mark. In an attempt to isolate the particular marking categories (variables) that were most often found to be missing or not observed on an officer's report, the distribution of marks for each of the marking categories (over all reports for all ranks) was examined. For each of the 23 marking categories, Table II shows the percentage of fitness reports on file on which that particular category was marked not observed or simply left blank. Constructing a variable set composed of variables for which few reports had missing values, and using this set in a discriminant analysis, would mean that close to 100% of the officer records would be used in the analysis. This in itself was reason enough to study the effectiveness of discriminant functions based on such sets. Therefore, for each of the three original sets (BN, WB, and GB), an additional subset was constructed composed of only those marking categories displaying less than a 30% missing value rate. These reduced variable sets carry the name of the original set from which they were formed, but with a suffixed -R (e.g., BNR).

In summary, the 12 variable sets to be used in the construction of the discriminant functions to be analyzed in the following sections are listed--along with their descriptions--in Table III.

TABLE II

PERCENTAGE OF NOT OBSERVED OR MISSING MARKS FOR EACH CATEGORY

Category	% Not Observed	Category	% Not Observed
*13a	22.4	*14f	20.3
13b	76.5	*14g	21.1
*13c	28.5	14h	91.7
13d	39.9	*14i	21.6
13e	33.6	*14j	23.1
13f	44.9	*14k	20.5
13g	81.9	*14l	20.4
14a	66.9	14m	31.8
*14b	19.3	14n	48.7
*14c	19.4	*15a	22.2
*14d	19.5	*16	4.3
*14e	19.6		

*used in reduced variable set

TABLE III

DESCRIPTION OF VARIABLE SETS

Variable Set	Description
BN	The mark on each category in this set is the mean over all reports for that mark
BNW	The variable set BN reduced by the forward selection technique
BNR	Similar to set BN, but only those marking categories asterisked in Table II are included
BNRW	The variable set BNR reduced by the forward selection technique
WB	The mark on each category in this set represents the length-weighted average over all reports for that mark
WBW	The set WB reduced by the forward selection technique
WBR	Similar to set WB, but only those marking categories asterisked in Table II are included
WBRW	The set WBR reduced by the forward selection technique
GB	Similar to set WB, but only those reports written while in an officer's present grade are considered
GBW	The variable set GB reduced by the forward selection technique

TABLE III

DESCRIPTION OF VARIABLE SETS

(Continued)

Variable Set	Description
GBR	Similar to set GB, but only those marking categories asterisked in Table II are included
GBRW	The set GBR reduced by the forward selection technique

B. METHOD OF ANALYSIS

The problem of choosing that particular discriminant function yielding discriminant scores which best serve as performance indices is in fact a problem of choosing that set of discriminating variables which, when used in the discriminant analysis, yields the best discriminant function. Therefore, within each grade, discriminant functions computed using the 12 different variable sets will be subjected to several criteria with the intent of finding the optimum variable set for that grade, for this year's data.

The criteria to be used will be as follows:

The first criterion will be the quantitative p criterion described in Chapter II and reiterated here. Given a discriminant function computed for a particular grade using one of the 12 variable sets, discriminant scores will be determined for the officers of that grade, whose records were entered in the analysis, using the computed discriminant function. These officers will then be ordered by their discriminant score from highest to lowest. If, among these same officers, the actual number promoted is, say, n, then the proportion of promoted officers from the first n officers on the ordered list will be the criterion measure p.

The second criterion will be the proportion of officers of a particular grade whose variable sets were actually

entered into the analysis to obtain the discriminant function. This measure would also represent the proportion of officers of a particular grade for whom discriminant scores could be computed using that discriminant function.

It is recognized that there may exist other more elegant, perhaps more effective, criteria by which to measure differences among candidates. However, it must also be recognized that the two criteria mentioned must at least be included in any list of necessary measures. At any rate, the chapter will conclude by demonstrating that the "best" discriminant functions--those functions whose criterion measures were most satisfactory for each particular grade--do indeed perform admirably in ranking the officers by quality, thus lending support to the two above-mentioned criteria as effectiveness measures and to the discriminant analysis technique as a method of constructing a performance index.

(Throughout the next few sections, bear in mind that the records of captains would appear before a promotion board to major and the records of majors before a promotion board to Lieutenant Colonel, etc. Whenever results are presented which involve promotion board action--for instance, the determination of an optimum variable set--the results will be under the name of the board. But whenever the results are to be applied to a particular grade of officer--the methodology of computing discriminant scores for the

rank of Captain, for example--the results will be presented by rank).

C. RESULTS

1. Majors' Board

The 9175 fitness reports of the 324 captains who appeared before the FY81 majors' promotion board were analyzed. Of the 324 captains in the analysis, 242 had in fact been selected for promotion for a rate of 74.7%. The results of the analysis for this board appear in Table IV.

Many of the results found for the majors' board could have been predicted beforehand. For example, the officers in the grade of captain had, on the average, fewer reports on file than officers of the two higher grades. As a result, the incidence of missing values on the composite marks in the complete (non-reduced) variable sets would be substantial. The results substantiate this. As seen in Table IV, the discriminant function based on the variable set BN gives, by far, the best measure of p. Yet only 37% of the officers had records which were admissible into the analysis. Of those variable sets reduced by the forward selection procedure, BNW gives the best combined performance on both criteria, yet still permits only 81% of the records to enter the analysis.

TABLE IV

MAJORS' BOARD RESULTS

Variable Set Entered into Analysis	% Exhibiting NO Missing Values	*p
BN	37	.958
BNW	81	.902
BNR	100	.901
BNRW	100	.909
WB	36	.937
WBW	81	.887
WBR	100	.905
WBRW	100	.905
GB	12	.914
GBW	45	.786
GBR	100	.896
GBRW	100	.892

*Actual promotion opportunity: .747

The results for the variable sets that were reduced by restricting the marking categories based on percentage of missing values are also shown in Table IV. As it turns out, 100% of the officers' records were admissible into the analysis for all six of these variable sets. Of these sets, BNRW has the best measure in the p criterion.

Summarizing, the fact that fully 100% of the captains' records were used in the analysis when the variable set BNRW was used, coupled with the fact that the p criterion measure for BNRW is bettered only by variable sets allowing less than 40% of the records into the analysis, suggests that the set BNRW ought to be used for this rank.

2. Lieutenant Colonels' Board

16,609 fitness reports for the 378 Majors appearing before the FY81 Lieutenant Colonels' promotion board were analyzed. Of these 378 Majors, 254 were selected for promotion for a rate of 67.2%. The results for this board are shown in Table V. The best overall performance on the p criterion was obtained with a complete (non-reduced) variable set, GB. Yet only 18% of the records were used in the analysis. Again, 100% of the officers' records were admissible in the case of the reduced sets (those sets whose categories were included by virtue of their

TABLE V

LT COLS' BOARD RESULTS

Variable Set Entered into Analysis	% Exhibiting NO Missing Values	*p
BN	79	.879
BNW	99	.881
BNR	100	.866
BNRW	100	.858
WB	79	.874
WBW	99	.866
WBR	100	.858
WBRW	100	.862
GB	18	.957
GBW	64	.817
GBR	100	.85
GBRW	100	.846

*Actual promotion opportunity: .672

missing value percentage). However, the set BNW, which allows fully 99% of the Majors' records to be used in the analysis, has a p criterion bettered only by the variable set GB which permits only 18% participation. Therefore, it seems that variable set BNW is the best for this particular grade.

3. Colonels' Board

The results of the analysis on the Colonels' board are shown in Table VI. A total of 16,530 reports for 265 officers were examined. Of these 265 officers, 143 were selected for a promotion rate of 54%. As was true for the Lieutenant Colonel's board, the GB variable set has the best measure on the p criterion, but a mere 19% of the records were entered into the analysis. The differences between the remaining sets, however, are less distinctive. Set WBW has a full 99% of the records entering into the analysis, yet its p measure is not as good as the measure for set BN. In other words, there is no clear-cut choice to be made.

D. CONCLUSIONS

Especially in light of the results on the Colonels' board, it's difficult to rationalize making a decision on the best variable set for a particular rank without taking into account the results on the other ranks. Specifically, the results on both the Majors' and Lieutenant Colonels' boards suggest that variable sets whose category marks

TABLE VI

COLONELS' BOARD RESULTS

Variable Set Entered into Analysis	% Exhibiting NO Missing Values	*p
BN	87	.846
BNW	92	.821
BNR	100	.769
BNRW	100	.776
WB	87	.838
WBW	99	.809
WBR	100	.755
WBRW	100	.762
GB	19	.892
GBW	100	.671
GBR	100	.769
GBRW	100	.769

*Actual promotion opportunity: .54

represent the simple averages over all reports of the marks for that category--or subsets of this particular set--yield the best discriminant functions to be used in computing discriminant scores. This is fortunate since this variable set is, computationally, the least involved. So it would seem reasonable to add weight to the variable set BN in the Colonels' board analysis and, in so doing, choose it as the variable set to be used.

Summarizing, then, the results on the three boards suggest the following variables sets be used in the respective grades:

- For the grade of captain ~ set BNRW
- For the grade of major - set BNW
- For the grade of Lieutenant Colonel - set BN

The fact that the "completeness" of the optimum variable set for each rank (i.e., BN has more variables--is more complete--than BNW) must increase with an increase in rank is interesting. This point, among others, will be addressed in the following chapter.

E. AN HYPOTHESIZED MODEL

From the beginning, the proposition has been made and supported, at least in theory, that a ranking of officers by discriminant score would approximate a ranking by promotability. And, again, promotability is interpreted as quality for reasons that were outlined in Chapter II. In

other words, given a list of officers ranked by discriminant score, the promotion probability of the officer who occupies the i-th position on the list should be greater than or equal to the promotion probability of the officer who occupies the i-th + 1st position, if the proposition is true. In more specific terms, assign a sequence number to each of the officers in the ordered list. The first officer on the list--the one with the highest discriminant score--would be assigned a sequence number of 1. The second highest on the list would have a sequence number of 2, and so forth. Then divide the officers in this ordered list into k sequential blocks with each block i ($i=1,\ldots,k$) having n_i officers. The first block would contain those officers with sequence numbers 1 through n_1, the second block would contain sequence numbers n_1+1 through n_1+n_2, and so forth. Count the number of officers in each group i who were promoted, pr_i. Then the promotion proportion for the i-th group would be pr_i/n_i. What the proposition would suggest in this set-up is that the promotion proportion of each of the k groups is a decreasing function of the relative placement of that group among the others. In other words, the promotion proportion of the i-th group, $pr_i/n_i = p_i$ should be greater than or equal to the proportion of the i-th + 1st group, $pr_{i+1}/n_{i+1} = p_{i+1}$.

A simple k-sample chi-squared test for similar proportions among the k groups would, if rejected, imply that there is indeed a difference in the proportions but would suggest nothing else.

Of interest in this case, however, is that, given there is a difference in group proportions, do those proportions vary in relation to, say, the mean sequence number of the groups?

Fleiss [Ref. 7] outlines a statistical technique designed to investigate this last question. The technique is, in effect, a variation on the simple chi-square test for equal proportions. The hypothesis to be tested with this more detailed chi-square analysis, however, is that there is a significant tendency for the group proportions to vary with mean group sequence number. Different methods of analysis are called for depending on how the proportions are hypothesized to vary, but only the simplest kind of variation will be considered here--a linear one.

That is, let p_i again be the proportion of promoted officers in the i-th group and let x_i be the mean sequence number of the i-th group. The hypothesized model, then, giving the relationship between group proportion and group mean sequence number is:

$$p_i = \alpha + \beta x_i$$

where β, the slope of the line, indicates the amount of change in the proportion per unit change in sequence number and α, the intercept, indicates the proportion expected when $x = 0$.

The question may arise as to why a linear relationship is being hypothesized when in fact the variation of p_i with x_i may not be linear. What is of paramount interest is the decreasing nature of the relationship, i.e., p_i as a decreasing function of x_i. The actual shape of the function is of secondary importance. But, hopefully, two things are true:

- First, the relationship--if not exactly linear--is not too far removed from it, at least not far enough to result in rejection of the hypothesized model.
- Secondly, if in fact the hypothesized model is rejected, it will be as a result of the actual shape of the function, and not because the function is not indeed decreasing. Other evidence (graphs) will, hopefully, support this.

At any rate, a negative sign for the slope estimate, b, would indicate that the function is indeed decreasing. The significance of the "negativity" of this term can be tested, as will be seen.

The test proceeds as follows:

Let

n_i = number of officers in the i-th group

$$x_i \quad = \quad \text{mean sequence number of the ith group}$$

$$n. \quad = \quad \sum_{i=1}^{k} n_i$$

$$\bar{x} \quad = \quad \sum_{i=1}^{k} n_i x_i / n.$$

$$\bar{p} \quad = \quad \sum_{i=1}^{k} n_i p_i / n.$$

Then the slope term, β, is estimated by

$$b \quad = \quad \frac{\sum_{i=1}^{k} n_i p_i x_i \; - \; n.\bar{p}\,\bar{x}}{\sum_{i=1}^{k} n_i (x_i \; - \; \bar{x})^2}$$

and the intercept term, α, is estimated by

$$a \quad = \quad \bar{p} - b\,\bar{x}$$

Now, using the expression

$$\hat{p}_i \quad = \quad a + b\,x_i$$

or, equivalently,

$$\hat{p}_i \quad = \quad \bar{p} + b(x_i \; - \; \bar{x})$$

the expected proportion, $\hat{p}_i$, for each group can be calculated. The chi-square statistic is computed using

$$\chi^2 = \sum_{i=1}^{k} n_i(p_i - \hat{p}_i)/\overline{p}(1 - \overline{p})$$

and has k-2 degrees of freedom (where k is the number of groups).

The magnitude of this statistic will indicate the extent of the differences between the true proportions in each group (p_i) and the expected proportions, proportions that would be expected if the linear relationship was true. A small value of the chi-square statistic would tend to support the linear model, a large value to reject it.

Fleiss also shows that the following statistic,

$$\chi^2_{slope} = b^2 \sum_{i=1}^{k} n_i(x_i - \overline{x})^2/\overline{p}(1 - \overline{p})$$

is distributed chi-square with one degree of freedom and allows one to test the significance of the difference between the computed slope estimate, b, and 0.

The test was run using the discriminant functions determined for each of the three ranks. A discriminant score was computed for each officer of that particular rank. The officers were then ranked from top to bottom by discriminant score and

arranged into sequential groups. Of course, the determination of each group siz.., n_i, was dictated by the normal chi-square requirements such as fewer than 20% of the groups should have an expected number of promoted officers ($n_i\hat{p}_i$) less than 5, etc. The results were as follows:

1. <u>Majors' Board</u>

For the grade of captain, a graphical representation of the relationship between p_i and x_i is shown in Figure 4. Evident immediately is the general decreasing tendency of the group promotion proportions with an increase in group mean sequence number--a tendency that was expected. Also to be expected are the fluctuations in the proportions. The discriminant analysis technique, as has been mentioned many times before, takes into account only a small portion of the personal data that are used in the promotion board's decision. So a smooth curve--strictly decreasing--would certainly not be expected. Due to the high overall promotion opportunity for the captains appearing before the major's board (74.7%), a group promotion proportion close to or at 1.0 for the first few groups could also have been predicted. This is also an explained departure from linearity--but in no way detracts from the effectiveness of the technique for this grade.

The chi-square statistic for linearity is, for the captains' data, 10.25 which, with 4 degrees of freedom, is significant at the .04 level. The association with sequence number of the proportion of officers promoted is thus not

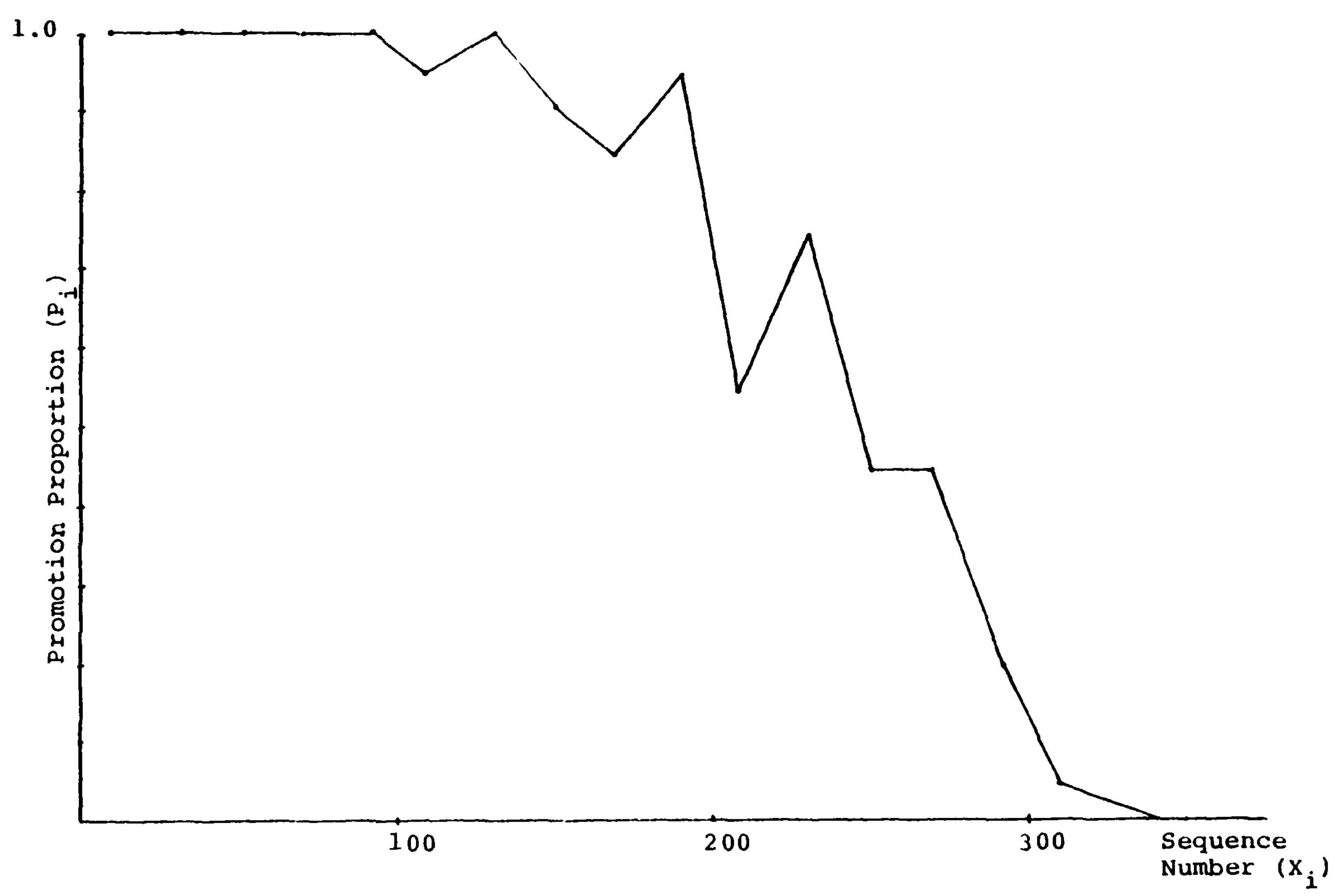


FIGURE 4. GRAPH OF SEQUENCE NUMBER VS PROMOTION PROPORTION (MAJORS' BOARD)

precisely a linear one, but the departures from linearity can be explained, and in fact are not severe. Further, the slope term, which is of course negative, is significantly different from 0.

2. Lieutenant Colonels' Board

Figure 5 depicts the relationship of p_i with x_i for the grade of major. The general decreasing tendency of the relationship is again evident, but the departures from linearity in this case are much less notable due in part to the lower promotion rate for this grade (67.2%).

Lending support to the graphical evidence are the results of the proportions test for the majors' data. The chi-square statistic was 1.805 which, with 6 degrees of freedom, was significant at greater than the .9 level. The slope term, b, was again negative and significantly different from 0.

3. The Colonels' Board

Again, the general decreasing relationship of group proportions with mean group sequence number is evident from Figure 6. Two other characteristics of the curve in Figure 3 could probably have been predicted. One, the relatively low overall promotion rate for this grade (54%) would result in a more rapid decrease in group promotion proportions with increase in sequence number in contrast to, say, the majors' board. Two, decisions on promotion to the grade of colonel involve a vast number of factors, few of them contained in

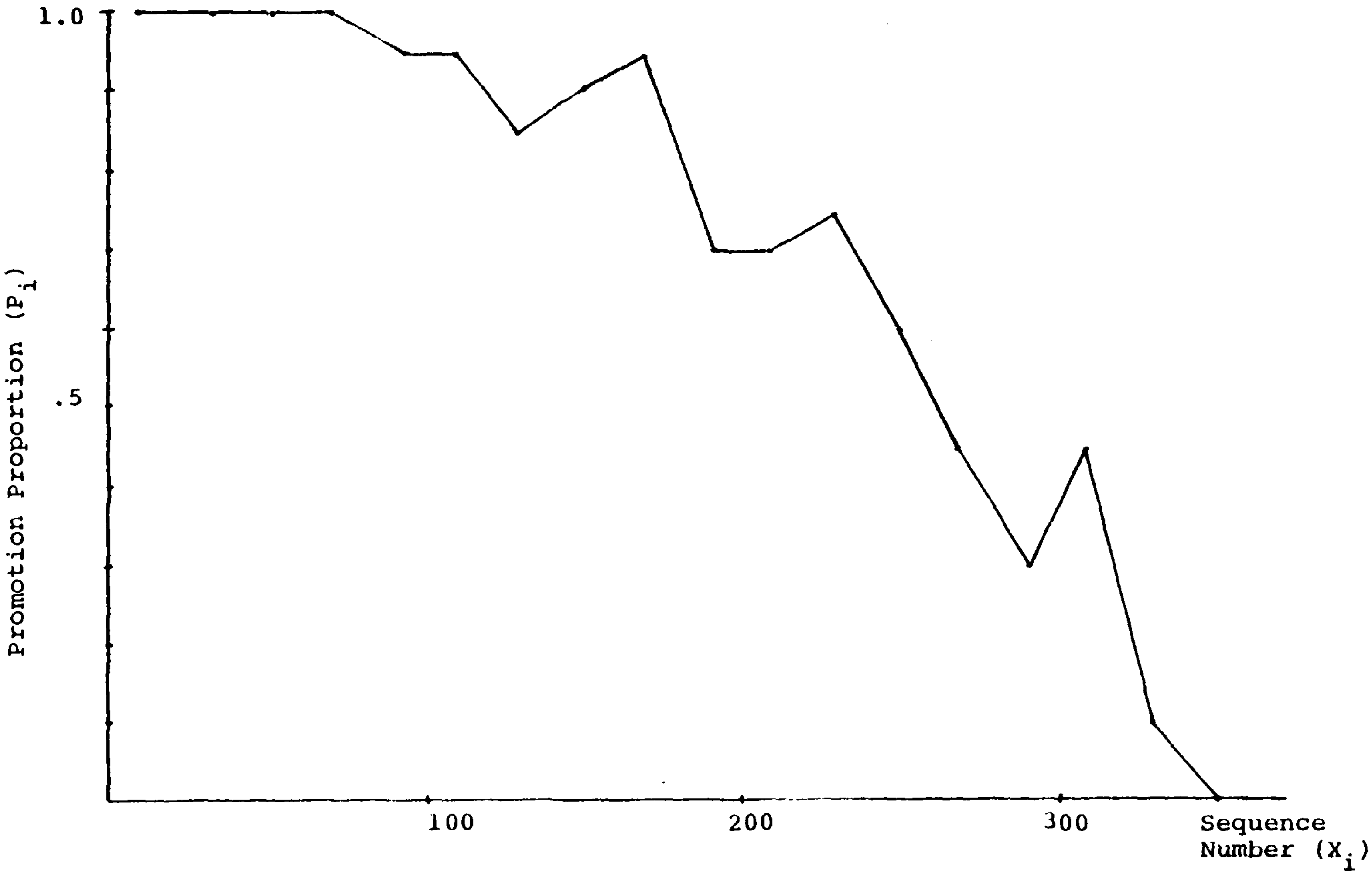


FIGURE 5. CRAPH OF SEQUENCE NUMBER VS PROMOTION PROPORTION (*LTCOLS' BOARD*)

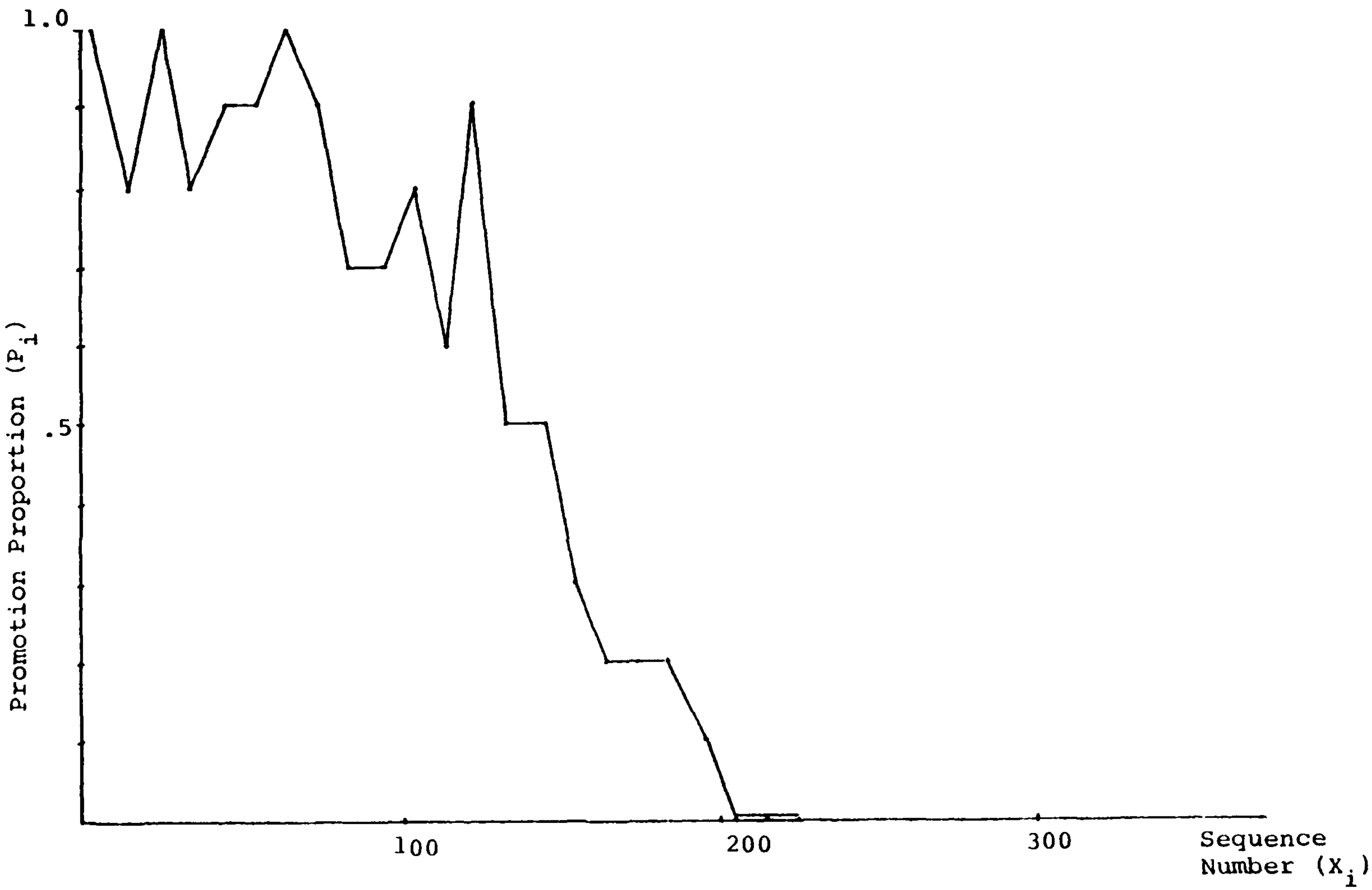


FIGURE 6. GRAPH OF SEQUENCE NUMBER VS PROMOTION PROPORTION (COLONELS' BOARD)

the 23 fitness report categories that, of necessity, were used by this analysis. As a result, the discriminant technique is less effective as a performance index. This is evidenced by the wide fluctuations in promotion proportions in the curve.

Still, the chi-square statistic for linearity for this data was 9.354 which, with 7 degrees of freedom, was significant at the .25 level. The slope term, again negative, was significantly different from 0. So the hypothesized linear model is reasonable even at this grade.

In summary, it can be said with a certain degree of confidence that the proposition concerning the discriminant technique as an effective method for constructing a performance index is supported. It has been shown that once a discriminant function has been computed for officers of a certain grade using the applicable variable set, and these officers are then placed in an ordered list ranked by discriminant score, the promotion probability of an officer decreases in (generally linear) relation to his sequence number on that list.

F. IMPLEMENTATION

The implications of the supported preposition are as follows:

Take any group of officers of a particular grade--not necessarily those officers whose records were used in the analysis. If discriminant scores were computed for those officers using the discriminant function determined for that

particular grade, and if these same officers were ranked on a list by discriminant score, then it's not unreasonable to claim that the promotion probability of an officer on that list (if he had conjecturally appeared before a board) would generally decrease with an increase in his sequence number on that list. And, as has been mentioned before, if one were to equate promotability with quality, then the ordered list ranks those officers by quality--a basic requirement of a performance index.

The analysis in this thesis has dealt exclusively with the three grades of Captain, Major, and Lieutenant Colonel. The analysis for the other grades, however, would be identical. The main thrust of the analysis, as was the case in this thesis, is the determination of the variable set to be used in the discriminant routine. Once that is accomplished, the computation of discriminant functions is faciliated by the widespread availability of discriminant analysis routines in commercial computer packages--in particular, the SPSS package installed on the system at HQMC, Washington, D.C.

In the case of the three grades used in the analysis, the implementation of the results of this study go as follows:

For an officer of the grade of Captain, Major, or Lieutenant Colonel, compute his applicable variable set according to the results contained in Section D. Compute his discriminant score by multiplying the composite mark for each of the categories in the variable set by the corresponding coefficient,

found in Table VII, and then adding the constant, also found in the table. In the case of an officer of a particular rank who doesn't have a complete variable set of category marks and for whom, therefore, a discriminant score cannot be computed, the following procedures are suggested. When the discriminant scores are being utilized as a means of distributing record books among the members of a promotion board, a simple random procedure (e.g., alphabetically) for distributing the books of those officers without computed discriminant scores may be utilized. Efforts at estimating discriminant scores for those officers would not be simple and probably not worthwhile simply because the vast majority of the books would have been distributed correctly, still ensuring each promotion board member a representative cross-section of the officers under consideration for promotion. When the discriminant scores are being utilized in regression studies (the use of discriminant scores in this context was addressed in Chapter II, Section B), consideration in the analysis of only those officers with computed scores would probably have little or no effect on results.

(The theoretical determination of the discriminant function, presented in Chapter III, didn't involve the addition of a constant in the computation of discriminant scores. However, when discriminant scores are determined using raw (non-standardized) values for the composite category marks, the introduction of the constant is necessary to achieve

TABLE VII

COMPUTATION OF DISCRIMINANT SCORES

(Rank of Captain)

Multiply the Composite Mark For	By
13a	.6547828
14b	-.5775795
14f	-.5628497
15a	3.593166
16	2.983457
Constant	-52.0217

(Rank of Major)

13d	.6675456
14a	.3491884
14c	.4424754
14d	-.7007447
14e	-.6392594
14f	.5063039
14g	-1.048672
14i	-.4696341
14n	1.75427
15a	3.173388
Constant	-33.74714

TABLE VII (Cont.)

(Rank of Lt Col)

Multiply the Composite Me For	By
13a	-.5688985
13b	-.4045532
13c	.6361058
13d	.8929157
13e	.5612866
13f	-.4781318
13g	.1587282
14a	.3361749
14b	-.1872222
14c	.06260754
14d	-1.598732
14e	-1.543235
14f	.3524519
14g	-.6931471
14h	.1894585
14i	-.7334108
14j	1.557007
14k	2.348909
14l	-.4552717
14m	-.6295893
14n	2.309024
15a	3.711399
16	-4.149739
Constant	-12.52781

scores identical to those gotten using standardized data. And, as has been mentioned, the coefficients to be applied to raw data also differ from those to be applied to standardized data.)

In a sense, a discriminant function, which was computed based on the results of a particular board, reflects the views of that board--in that the prejudices shown by board members in considering certain marking categories more than others have a distinct effect on the final discriminant function. It may be true that respectable results could be expected using one particular function from year to year. However, to ensure the "currency" of the functions--meaning that they are an accurate reflection of current board thinking, the functions ought to be recomputed for each grade following publication of the results of a board to select officers to the next higher grade.

V. THE DISCRIMINATORY POWER OF THE FITNESS REPORT CATEGORIES

A. INTRODUCTION

Observing the discriminant function coefficients listed at the end of the previous chapter, one might reasonably be interested in what factors contribute to a certain fitness report category having a larger coefficient than another. Indeed, concerned might be the word since, in computing discriminant scores, the larger a category's coefficient, the more weight the mark on that category has in determining an individual's performance index.

The fact that certain categories don't even appear in some of the discriminant functions is perhaps an even greater source of consternation, but rather easily explained--at least more easily than the explanation concerning the relative size of the coefficients of the categories included in the functions.

First, recall that certain of the categories were excluded outright on several of the proposed variable sets since a mark on those variables was missing on over 30% of the reports anyway.

The exclusion of certain other of the variables was a result of the mechanics of the forward selection process. Briefly, if the discriminatory "information" provided by a category was contained in another category--or even in a

combination of two or more categories--already selected for analysis, it wouldn't be included. Further details concerning such mechanics can be found in Eisenbeiss [Ref. 5]. High degrees of correlation between the categories would make this effect even more severe. In the extreme, a correlation coefficient of 1.0 between all categories would result in only one category in the analysis.

The correlations between several of the fitness report categories was in fact high, and this also interferes with the strict interpretation of a category's coefficient magnitude as the discriminatory power of that category--an interpretation espoused, with little or no explanation, by several authors. Most often, the interpretation goes as follows: using the standardized discriminant function coefficients, the more "discriminating" a variable is in distinguishing one group from another, the greater will be the magnitude of that variable's coefficient.

But what is meant by "discriminating" and how is its degree to be measured? If an answer were to be found, steps could be taken to make the more important categories more discriminating and the magnitude of their associated coefficients greater--thereby increasing the influence of those categories in determining discriminant scores. Intuitively, several possible answers come to mind.

First, it seems that certain statistics connected with the marking categories might have a correlation with

discrimination. For instance, the greater the variance or the range of marks in a category, the greater the possible distance between the group means on that category, and the larger, perhaps, the magnitude of its coefficient.

A second intuitive answer, and certainly a desired one, is that the actions of the board in relying more heavily on the marks in certain categories in making their promotion decision should have an effect on the magnitude of the discriminant coefficients.

Both possibilities will be briefly investigated in the following sections.

B. CATEGORY STATISTICS

Rigorous proofs of the existence of hypothesized correlations between variable parameters and discriminant coefficient magnitudes would be beyond the scope of this thesis--if indeed the correlations existed.

The tack taken here, however, will be a simple empirical examination of the relationships--if any--between certain statistics computed on the composite marks for each of the 23 fitness report categories and the discriminant coefficient computed on that same composite mark. This will be done for the data on each of the three grades of Captain, Major, and Lieutenant Colonel. The composite mark for each category will be the simple average over all reports on file for that category. In other words, the variable

set BN will be entered into the discriminant analysis for each grade. Even though BN was not the recommended variable set for either of the grades Captain or Major, subsets of BN in fact were. In addition, a variable set that included all the categories was needed. The statistics to be examined are the following:

- The average marks on each category for the promoted officers
- The same average mark for the not promoted officers
- The difference between the two means
- The combined (promoted and not promoted officer records taken together) range of marks on each category
- The combined sample variance of the marks in each category
- The F-ratio for each category mark. The F-ratio to be defined as the ratio of the mean squares due to group means and the pooled within-group sample variance, as defined in Chapter III, Section B.

These statistics are displayed for all three grades in Table VIII. The results are interesting in themselves, but, unfortunately, there doesn't seem to be any clear-cut relationship between coefficient magnitude and any of the statistics. The only possible exception is the association of coefficient magnitude with the F-ratio. For example, over all three grades, the largest discriminant coefficient

TABLE VIII

MARKING CATEGORY STATISTICS

RANK OF CAPTAIN

Composite Mark For	Magnitude of DF Coeff.	Prom Mean	Not Prom Mean	Difference	Combined Range	Combined Variance	F-Ratio
13a	.21404	8.263	7.729	.534	2.25	.165	50.52
13b	.12884	8.079	7.589	.49	5.0	.608	12.91
13c	.37466	7.728	7.279	.449	2.857	.233	14.84
13d	.06505	7.972	7.517	.455	2.5	.233	12.72
13e	.25895	8.064	7.647	.417	2.462	.207	19.22
13f	.00423	8.119	7.693	.426	2.6	.195	25.42
13g	.18906	8.025	7.584	.441	4.0	.519	15.01
14a	.0296	8.296	8.051	.245	2.667	.411	.626
14b	.26724	8.355	8.157	.198	2.182	.243	2.737
14c	.04081	8.34	8.019	.321	2.0	.157	16.96
14d	.04177	8.485	8.057	.428	1.909	.144	38.71
14e	.41430	8.564	8.309	.255	1.636	.099	18.62
14f	.15358	8.252	7.757	.495	2.56	.203	35.89

TABLE VIII

MARKING CATEGORY STATISTICS

(Continued)

RANK OF CAPTAIN

Composite Mark For	Magnitude of DF Coeff.	Prom Mean	Not Prom Mean	Difference	Combined Range	Combined Variance	F-Ratio
14g	.18244	7.86	7.355	.505	2.64	.201	26.17
14h	.08990	8.0	7.939	.061	4.0	1.116	.326
14i	.16738	7.919	7.552	.367	2.667	.244	15.8
14j	.08998	8.218	7.738	.48	2.56	.185	30.09
14k	.24225	8.848	8.739	.109	.923	.03	10.53
14l	.0073	8.339	7.985	.354	2.143	.184	19.01
14m	.07231	7.886	7.565	.315	2.346	.136	18.78
14n	.35551	8.607	8.184	.423	2.0	.126	52.95
15a	.84524	8.279	7.754	.525	2.115	.132	69.96
16	.30381	8.993	8.953	.04	.452	.002	18.43

TABLE VIII

MARKING CATEGORY STATISTICS

(Continued)

RANK OF MAJOR

Composite Mark For	Magnitude of DF Coeff.	Prom Mean	Not Prom Mean	Difference	Combined Range	Combined Variance	F-Ratio
13a	.07203	8.196	7.735	.461	2.146	.163	128.3
13b	.00631	7.865	7.355	.51	4.0	.533	37.73
13c	.08983	7.744	7.329	.415	2.522	.242	43.3
13d	.23649	8.036	7.57	.466	2.261	.187	107.3
13e	.00827	8.084	7.659	.425	2.267	.185	81.13
13f	.04952	8.066	7.667	.399	2.609	.204	70.7
13g	.0545	7.942	7.48	.462	6.0	.691	21.18
14a	.16826	8.223	7.785	.438	2.4	.289	45.21
14b	.0245	8.337	7.895	.442	2.632	.268	60.78
14c	.10338	8.315	7.909	.406	2.089	.169	85.73
14d	.25367	8.46	8.154	.306	2.044	.142	49.29

TABLE VIII

MARKING CATEGORY STATISTICS

(Continued)

RANK OF MAJOR

Composite Mark For	Magnitude of DF Coeff.	Prom Mean	Not Prom Mean	Difference	Combined Range	Combined Variance	F-Ratio
14e	.14941	8.364	8.14	.224	2.0	.118	26.94
14f	.18578	8.225	7.789	.436	2.444	.183	94.1
14g	.34559	7.87	7.408	.462	2.542	.195	87.8
14h	.02459	8.129	7.797	.332	4.0	.699	8.56
14i	.29467	7.964	7.533	.431	2.487	.259	58.24
14j	.18752	8.199	7.719	.48	2.318	.178	115.5
14k	.01563	8.811	8.772	.039	1.023	.028	15.51
14l	.11352	8.251	7.947	.304	2.318	.149	43.53
14m	.02578	7.942	7.614	.328	2.309	.142	56.61
14n	.46646	8.764	8.355	.409	2.5	.115	151.3
15a	.84849	8.272	7.789	.483	2.181	.135	176.6
16	.02508	8.978	8.944	.034	.39	.002	25.82

TABLE VIII

MARKING CATEGORY STATISTICS

(Continued)

RANK OF LT COL

Composite Mark For	Magnitude of DF Coeff.	Prom Mean	Not Prom Mean	Difference	Combined Range	Combined Variance	F-Ratio
13a	.1912	7.9957	7.6729	.3228	2.098	.160	58.85
13b	.246	7.6787	7.3834	.2953	3.636	.458	16.17
13c	.26849	7.6610	7.3810	.28	2.56	.214	32.81
13d	.31888	8.0153	7.6767	.3386	2.1	.173	49.36
13e	.20379	7.9623	7.6474	.3149	2.196	.185	44.75
13f	.18629	7.9354	7.5811	.3543	2.498	.195	43.43
13g	.16624	7.6204	7.0827	.5377	8.0	1.165	14.09
14a	.14775	8.1405	7.7875	.353	2.6	.258	42.77
14b	.07750	8.2709	7.9482	.3227	2.215	.203	33.54
14c	.02281	8.2299	7.9143	.3156	2.191	.168	44.5
14d	.55118	8.3406	8.1214	.2192	1.815	.139	29.33
14e	.56281	8.1632	8.0226	.1406	1.959	.147	8.966

TABLE VIII

MARKINC CATEGORY STATISTICS

(Continued)

RANK OF LT COL

Composite Mark For	Magnitude of DF Coeff.	Prom Mean	Not Prom Mean	Difference	Combined Range	Combined Variance	F-Ratio
14f	.13352	8.0801	7.7737	.3064	2.277	.179	42.81
14g	.25764	7.7730	7.4675	.3055	2.334	.177	43.58
14h	.1212	8.4505	8.1177	.3328	4.0	.448	13.46
14i	.30583	7.9535	7.6052	.3483	2.481	.216	37.22
14j	.52008	8.1103	7.7547	.3556	2.128	.162	65.41
14k	.38437	8.7372	8.6511	.0861	.966	.03	23.28
14l	.16913	8.1130	7.9443	.1687	2.449	.160	12.79
14m	.21271	7.8264	7.6451	.1813	2.205	.139	19.24
14n	.47734	8.8509	8.5996	.2513	1.1	.062	77.54
15a	1.11454	8.1260	7.8043	.3217	1.927	.130	70.75
16	.22205	8.9433	8.9121	.0312	.25	.003	19.16

and the largest F-ratio are associated with the same category. But this certainly isn't a surprising or a particularly useful result. For one thing, the statistics that determine the F-ratio are a by-product of board action. The F-ratio is not associated with the category itself since a different grouping of officers into the promoted and not promoted groups would have resulted, perhaps, in a different mean square due to group means on the category and thence a different F-ratio. Secondly, the computed discriminant coefficients were produced in an attempt to find a linear combination of the original category marks that would give maximum separation of group means on that linear combination relative to within-group variances on that combination. In a sense, this same group mean difference relative to within-group variation is what the F-ratio also expresses. So it would seem reasonable to expect a category having a large F-statistic to also contribute heavily to the above-mentioned linear combination--i.e., have a large discriminant coefficient.

However, as can be seen from Table VIII, those statistics which are independent of board action, namely range and variance, seem to have no obvious correlation with coefficient magnitude.

C. POSSIBLE EFFECTS OF BOARD BIAS

The second possible explanation for the different coefficient magnitudes was the effect of board members' predispositions in weighting certain categories more heavily when

rendering their promotion decisions. That this may be so is suggested to a degree by the values in Table VIII. Traditionally, such categories as "general value to the service" (item 15a), "leadership" (item 14j), and "growth potential" (item 14n) have been among the most important and influential on the fitness report. The coefficients for each of these categories are relatively large over all these grades. At the same time, one would expect the relative importance of certain of the categories to change, depending on the particular board. For instance, "handling of officers" (item 13d) and "economy of management" (item 14m) are important measures of a Lieutenant Colonel's quality and perhaps to a lesser extent, a measure of a Captain's This supposition is also supported by values in Table VIII.

The effects of board bias on discriminant coefficients can also be demonstrated using a simple mathematical model. The analysis will exploit the following fact, covered in detail by Tatsuoka [Ref. 1].

To each officer's standardized p-vector of category marks, attach a dependent, dichotomous variable (such as a variable taking values 1 or 0) representing the officer's group. For instance, let 1 mean the officer was promoted and 0 mean not promoted. In a regression of this dichotomous variable on the p-vector of category marks, the computed regression weights will differ only by a multiplicative constant from the discriminant coefficients determined

by entering the p-vectors of category marks into a discriminant analysis.

So, to show that board bias affects discriminant coefficients, it will suffice to show that the bias would effect the regression weights.

Suppose that all officers before a particular board were characterized by marks on only two categories, A and B. That is, the ith officer's 2-vector of marks would be (a_i, b_i). Suppose further that the board decided to base their promotion decisions only on the mark for category A, completely ignoring the B category mark. (Admittedly, this is an example of extreme, most improbable bias yet the example will serve to illustrate a point.) In addition, the marks on each of the categories represent standardized scores so that

$$\sum a_i = 0 \qquad \text{and} \qquad \sum b_i = 0$$

For simplicity, assume that the board considered only four officers, of whom two were promoted, and that all the promoted officers had a score of 1 for category A and the not promoted officers a score of -1. The scores on category B can take on any values, as long as they sum to 0.

Attaching the dependent variable (1,0) to the applicable 2-vectors of category marks, the data are:

	Dep variable	Mark for A	Mark for B
officer 1	1	1	b_1
2	1	1	b_2

3	0	-1	b_3
4	0	-1	b_4

and the regression model becomes

$$Y = X\beta + e$$

where β is the (2x1) matrix of regression weights, X is the (4x2) matrix of marks for A and B, Y is the (4x1) matrix of dependent variables (1,0), and e is the (4x1) matrix of error terms.

It now remains to show that the board's predisposition in favoring one category over the other will result in a regression weight for category A (β_A) larger than that for category B (β_B).

The expression for the estimated regression weights $\hat{\beta}' = (\hat{\beta}_A, \hat{\beta}_B)$ is

$$\hat{\beta} = (X'X)^{-1} (X'Y)$$ (where ' indicates transpose and $^{-1}$ indicates inverse).

Solving for each of the elements of $\hat{\beta}$, it can be shown that

$$\det(X'X)\hat{\beta}_A > 0$$

and

$$\det(X'X)\hat{\beta}_B = 0.$$

The results prove that the predisposition of the board in considering only the one category--A--in making the promotion decision had the effect of making the regression coefficient for that category larger. Therefore, the discriminant coefficient for the category A, if the four 2-vectors of category marks had been entered into a discriminant analysis, would also have been larger. It would also follow that the dependent variable score would be unaffected by B values, since a change in B is multiplied by 0, and so the same would be true for the discriminant score.

In the case where there are more than two categories, or where the consideration of categories is not so one-sided, the mathematics of the model are not nearly so tractable, but the results would be similar--i.e., the more consideration given a particular category when the promotion decision is made, the larger the discriminant function coefficient for that category.

D. CONCLUSION

The conclusion to be drawn from the analysis above is particularly appealing and lends considerable credibility to the discriminant analysis technique as a method for constructing a performance index.

There has been widespread and well-founded concern throughout the Marine Corps regarding the lack of distribution in the category marks on the fitness report. With few exceptions, all of the marks on a typical report are distributed exclusively

over the excellent and outstanding blocks. A solution to the problem will have to soon be found but the options will surely all take time to implement.

In the meantime, the discriminant analysis technique has shown its robustness even with this lack of distribution. Indeed, the range and variance of marks on a category seems empirically to have no effect on the coefficient for that category in the discriminant function.

What does have an effect is the inclination of the promotion board to rely more heavily on particular categories in rendering its promotion decision. In other words, a performance index determined by the discriminant function technique will generally reflect the same standards the promotion board used in judging the quality of the officers it considered.

VI. SUGGESTIONS FOR FURTHER STUDY

As has been mentioned, the main thrust of the analysis in Chapter IV was in the choice of variable set to be entered into the discriminant analysis. Although twelve sets were considered, the list of others that merit study is extensive. For instance, a set of composite category marks wherein a weight assigned to a particular report's mark is a decreasing function of age of the report (for example, weight = 1/(current date - report date)) might prove interesting. One of the important considerations in devising the different variable sets, however, is the logic of the weighting scheme employed. In other words, the scheme should be one that a promotion board member might be reasonably expected to employ when considering all the reports of a particular officer. With this in mind, the list of possible sets is probably not that extensive.

Another area that might benefit from further study is the question of measures of effectiveness for the different variable sets. Those employed in this thesis were characterized by their simplicity and intuitive appeal. Perhaps others could be devised, however, which would more rigorously examine proposed variable sets.

Still another area of interest would be the changes in discriminant weights over time. In other words, given a particular grade, would the discriminant weights for each

category remain generally constant when recomputed from year-to-year or would they change?

The ultimate test of the effectiveness of the discriminant technique would, of course, come when the discriminant scores computed for a particular promotion board were validated with the data from the subsequent board. Specifically, if one were to rank the officers appearing before the next promotion board by discriminant scores based on the discriminant function computed from this last board, would the results be supportive? That is, would the ranking by discriminant score again be generally a ranking by promotability, based on the subsequent board's promotion decisions? It is important that this be the case so the question would certainly be one worth investigating.

The shift in category discriminant weights with the rank of the board was suggested and investigated to some extent in the previous chapter. Further study might reveal information suggesting, perhaps, different report formats for different grades.

Finally, Amick [Ref. 8] suggests a method for determining the discriminant function which would exploit the high degree of correlation among several of the category marks. The method involves employing the Factor Analysis technique to first reduce the dimension of the variable set and then to enter the resultant factors into a discriminant routine. The reference outlines the advantages and disadvantages of such

a procedure. Its application to the performance index problem is evident and the method certainly bears study.

LIST OF REFERENCES

1. Tatsuoka, M.M., Multivariate Analysis, pp. 157-190, Wiley, 1971.

2. Rulon, P.J., "Distinctions between Discriminant Analysis and Regression Analysis and a Geometric Interpretation of the Discriminant Function", Harvard Educational Review, v. 21, pp. 80-90, 1951.

3. Eisenbeis, R.A. and Avery, R.B., Discriminant Analysis and Classification Procedures, p. 37, Lexington Books, 1972.

4. Morrison, D.G., "On the Interpretation of Discriminant Analysis", Journal of Marketing Research, pp. 156-163, May, 1969.

5. Eisenbeis, R.A. and Avery, R.B., op. cit., pp. 67-85.

6. Nie, N.H., and others, SPSS: Statistical Package for the Social Sciences, pp. 434-462, McGraw-Hill, 1975.

7. Fleiss, J.L., Statistical Methods for Rates and Proportions, pp. 96-99, Wiley and Sons, 1973.

8. Amick, D.J. and Walberg, H.J., Introductory Multivariate Analysis, pp. 244-246, McCutcheon Publishing Corporation, 1975.

BIBLIOGRAPHY

Amick, D.J. and Walberg, H.J., Introductory Multivariate Analysis, McCutcheon Publishing Corporation, 1975.

Brown, G.W., "Discriminant Functions", The Annals of Mathematical Statistics, v. 18, pp. 514-528, December, 1947.

Cooley, W.W. and Lohnes, P.R., Multivariate Data Analysis, Wiley, 1971.

Eisenbeis, R.A. and Avery, R.B., Discriminant Analysis and Classification Procedures, Lexington Books, 1972.

Fleiss, J.L., Statistical Methods for Rates and Proportions, John Wiley and Sons, 1973.

Garrett, H.E., "The Discriminant Function and its Use in Psychology", Psychometrika, v. 8, pp. 65-79, June, 1943.

Millard, C.A., A Factor Analysis of the Behavioral Dimensions of Marine Corps Performance Evaluation, Master's Thesis, Naval Postgraduate School, Monterey, Ca., 1976.

Morrison, D.G., "On the Interpretation of Discriminant Analysis", Journal of Marketing Research, pp. 156-163, May, 1969.

Nie, N.H., and others, SPSS: Statistical Package for the Social Sciences, McGraw-Hill, 1975.

Rulon, P.J., "Distinctions between Discriminant and Regression Analysis and a Geometric Interpretation of the Discriminant Function", Harvard Educational Review, v. 21, pp. 80-90, 1951.

Tatsuoka, M.M., Multivariate Analysis, Wiley, 1971.

Travers, R.M., "The Use of a Discriminant Function in the Treatment of Psychological Group Differences", Psychometrika, v. 4, pp. 25-32, March, 1939.

Van deGeer, J.P., Introduction to Multivariate Analysis for the Social Sciences, W.H. Freeman and Co., 1971.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 55 Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
4.	Professor H.J. Larson, Code 55La Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
5.	Professor R.S. Elster, Code 55Ea Department of Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
6.	Professor W.F. Moroney, Code 55Mp Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
7.	Commandant of the Marine Corps, Code RDS Headquarters, United States Marine Corps Washington, D.C. 20380	1
8.	Commandant of the Marine Corps, Code MPI-20 Headquarters, United States Marine Corps Washington, D.C. 20380	1
9.	Captain W.J. Haffey, USMC 142 Ribaut Rd. Beaufort, South Carolina 29902	1
10.	Deputy Chief of Naval Operations (MPT) Attn: OP-13 Arlington Annex Washington, D.C. 20370	1

END